

#30

1635

RAW SEQUENCE LISTING                      DATE: 11/29/2000  
PATENT APPLICATION:    US/09/446,402              TIME: 13:41:09

Input Set : A:\Lucyl.app  
Output Set: N:\CRF3\11292000\I446402.raw

## SEQUENCE LISTING

4 (1) GENERAL INFORMATION:  
6     (i) APPLICANT: Black Jr., Charles A.  
8     (ii) TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ACTIVATING  
9    GENES OF INTEREST  
11    (iii) NUMBER OF SEQUENCES: 16  
13    (iv) CORRESPONDENCE ADDRESS:  
14        (A) ADDRESSEE: W. Murray Spruill  
15        (B) STREET: 3605 Glenwood Ave. Suite 310  
16        (C) CITY: Raleigh  
17        (D) STATE: NC  
18        (E) COUNTRY: US  
19        (F) ZIP: 27622  
21    (v) COMPUTER READABLE FORM:  
22        (A) MEDIUM TYPE: Floppy disk  
23        (B) COMPUTER: IBM PC compatible  
24        (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
27    (vi) CURRENT APPLICATION DATA:  
C--> 28        (A) APPLICATION NUMBER: US/09/446,402  
C--> 29        (B) FILING DATE: 20-Dec-1999  
30        (C) CLASSIFICATION:  
32    (viii) ATTORNEY/AGENT INFORMATION:  
33        (A) NAME: Spruill, W. Murray  
34        (B) REGISTRATION NUMBER: 32,943  
35        (C) REFERENCE/DOCKET NUMBER: 5722-2  
37    (ix) TELECOMMUNICATION INFORMATION:  
38        (A) TELEPHONE: 919 420 2202  
39        (B) TELEFAX: 919 881 3175  
42 (2) INFORMATION FOR SEQ ID NO: 1:  
44     (i) SEQUENCE CHARACTERISTICS:  
45        (A) LENGTH: 4279 base pairs  
46        (B) TYPE: nucleic acid  
47        (C) STRANDEDNESS: single  
48        (D) TOPOLOGY: linear  
50     (ii) MOLECULE TYPE: other nucleic acid  
51        (A) DESCRIPTION: /desc = "Recombinant molecule  
W--> 52    (Multiple Cloning Site/Kozack sequence/LacZ gene)"  
55     (ix) FEATURE:  
56        (A) NAME/KEY: misc\_feature  
57        (B) LOCATION: 1..64  
58        (D) OTHER INFORMATION: /product= "Multiple Cloning Site"  
60     (ix) FEATURE:  
61        (A) NAME/KEY: misc\_feature  
62        (B) LOCATION: 65..79  
63        (D) OTHER INFORMATION: /function= "Consensus sequence of  
64 translation initiation"

ENTERED

## RAW SEQUENCE LISTING

DATE: 11/29/2000

PATENT APPLICATION: US/09/446,402

TIME: 13:41:09

Input Set : A:\Lucyl.app

Output Set: N:\CRF3\11292000\I446402.raw

```

65 /product= "Kozack sequence"
66 (ix) FEATURE:
67 (A) NAME/KEY: prim_transcript
68 (B) LOCATION: 80..4279
69 (D) OTHER INFORMATION: /gene= "LacZ"
70 /standard_name= "Beta galactosidase"
71 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
72 TTAATACGAC TCACTATAGG CTAGCCTCGA GAATTCACGC GTGGTACCTC TAGAGTCGAC 60
73 CCGGGCCGCG CCCACCATGG CGCAGCACCA TGGCCTGAAA TAACCTCTGA AAGAGGAACT 120
74 TGGTTAGGTA CCTTCTGAGG CGGAAAGAAC CAGCTGTGGA ATGTGTGTCA GTTAGGGTGT 180
75 GGAAAGTCCC CAGGCTCCCC AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA 240
76 GCAACCAGGT GTGAAAGTCC CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT 300
77 CPCAATTAGT CAGCAACCAT AGTCCCGGCC CTAACCTCCG CCATCCCGCC CCTAACTCCG 360
78 CCCAGTCCG CCATTCTCCG GCCCATGGC TGACTAATTT TTTTATTTA TGCAGAGGCC 420
79 GAGGCGCGCT CGGCCTCTGA GCTATTCAG AAGTAGTGAG GAGGCTTTTT TGGAGGCCCTA 480
80 GGCTTTTGCA AAAAGCTTGG GATCTCTATA ATCTCGCGCA ACCTATTTTC CCTCGAACA 540
81 CTTTTTAAGC CGTAGATAAA CAGGCTGGGA CACTTCACAT GAGCGAAAAA TACATCGTCA 600
82 CTTGGGACAT GTTGCAGATC CATGCACGTA AACTCGCAAG CCGACTGATG CCTTCTGAAC 660
83 AATGGAAGG CATTATTGCC GTAAGCCGTG GCGGTCTGGT ACCGCTGGGT GAAGACCAGA 720
84 AACAGCACCT CGAAGTGGG CGCGATATTG CCCAGCGTTT CAACGCGCTG TATGGCGAGA 780
85 TCGATCCCCG CTGTTTACAA CGTCGTGACT GGGAAAAACC TGGCGTTACC CAACTTAATC 840
86 GCCTTGCAGC ACATCCCCCT TTCGCCAGCT GGCCTAATAG CGAAGAGGCC CGCACCAGTC 900
87 GCCCTTTCCA ACAGTTGGCG AGCCTGAATG GCGAATGGCG CTTTGCCTGG TTTCCGGCAC 960
88 CAGAAGCGGT GCCGAAAGC TGGCTGGAGT GCGATCTTCC TGAGGCCGAT ACTGTCTGTC 1020
89 TCCCTCAAAA CTGGCAGATG CACGGTTACG ATCGGCCCAT CTACACCAAC GTAACTTATC 1080
90 CCATTACGGT CAATCCGCGG TTTGTTCCCA CGGAGAATCC GACGGGTTGT TACTCGTCA 1140
91 CATTTAATGT TGATGAAAGC TGGCTACAGG AAGGCCAGAC GCGAATTATT TTTGATGGCG 1200
92 TTAACCTCGC GTTTCATCTG TGGTGCAACG GCGCTGGGT CCGTTACGGC CAGGACAGTC 1260
93 GTTTGCCGTC TGAATTTGAC CTGAGCGCAT TTTTACGCGC CGGAGAAAAC CGCCTCGCGG 1320
94 TGATGGTGCT GCGTTGGAGT GACGGCAGTT ATCTGGAAGA TCAGGATATG TGGCGGATGA 1380
95 GCGGCATTTT CCGTGACGTC TCGTTGCTGC ATAAACCGAC TACACAAATC AGCGATTTC 1440
96 ATGTTGCCAC TCGCTTTAAT GATGATTTCG GCGCGCTGT ACTGGAGGCT GAAGTTTACA 1500
97 TGTGCGGCGA GTTGGTGGG TACCTACGGG TAACAGTTTC TTTATGGCAG GGTGAAACGC 1560
98 AGGTGCGCCG CGGCACCGCG CCTTTCGCGC GTGAAATTAT CGATGAGCGT GGTGGTTATG 1620
99 CCGATCGCGT CACACTACGT CTGAACGTCG AAAACCCGAA ACTGTGGAGC GCGGAAATCC 1680
100 CGAATCTCTA TCGTGCGGTG GTTGAAGTGC ACACCGCCGA CGGCACGCTG ATTGAAGCAG 1740
101 AAGCCTGCGA TGTGCGTTTC CGCGAGGTGC GGATTGAAA TGGTCTGCTG CTGCTGAACG 1800
102 GCAAGCCGTT GCTGATTCTG GCGTTAACC GTCACGAGCA TCATCCTCTG CATGGTCAGG 1860
103 TCATGGATGA GCAGACGATG GTGCAGGATA TCCTGCTGAT GAAGCAGAAC AACTTTAACG 1920
104 CCGTGGCGTG TTGCAATTAT CCGAACCATC CGCTGTGCTA CACGCTGTGC GACCGCTACG 1980
105 GCCTGTATGT GGTGGATGAA GCCAATATTG AAACCCACGG CATGGTGCCA ATGAATCGTC 2040
106 TGACCGATGA TCCGCGCTGG CTACCGCGCA TGAGCGAAGC CGTAACGCGA ATGGTGCAGC 2100
107 GCGATCGTAA TCACCCGAGT GTGATCATCT GGTGCTGGG GAATGAATCA GGCCACGGCG 2160
108 CTAATCACGA CGCGCTGTAT CGCTGGATCA AATCTGTCTG TCCTTCCCGC CCGGTGCAGT 2220
109 ATGAAGGCGG CGGAGCCGAC ACCACGGCCA CCGATATTAT TTGCCCGATG TACGCGCGCG 2280
110 TGGATGAAGA CCAGCCCTTC CCGGCTGTGC CGAATGGTC CATCAAAAAA TGGCTTTCGC 2340
111 TACCTGGAGA GACGCGCCCG CTGATCCTTT GCGAATACGC CCACGCGATG GGTAACAGTC 2400
112 TTGGCGGTTT CGCTAAATAC TGGCAGGCGT TTCGTCAGTA TCCCGGTTTA CAGGGCGGCT 2460
113 TCGTCTGGGA CTGGGTGGAT CAGTCGCTGA TTAAATATGA TGAAACGGC AACCCGTGGT 2520

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/446,402  
 DATE: 11/29/2000  
 TIME: 13:41:09

Input Set : A:\Lucyl.app  
 Output Set: N:\CRF3\11292000\I446402.raw

```

160 CGGCTTACGG CGGTGATTTT GCGGATACGC CGAACGATCG CCAGTTCCTGT ATGAACGGTC 2580
162 TGGTCTTTGC CGACCGCAGC CCGCATCCAG CGCTGACGGA AGCAAAACAC CAGCAGCAGT 2640
164 TTTTCCAGTT CCGTTTATCC GGGCAAACCA TCGAAGTGAC CAGCGAATAC CTGTTCCGTC 2700
166 ATAGCGATAA CGAGCTCCTG CACTGGATGG TGGCGCTGGA TGGTAAGCCG CTGGCAAGCG 2760
168 GTGAAGTGCC TCTGGATGTC GCTCCACAAG GTAAACAGTT GATTGAAGTG CCTGAACTAC 2820
170 CGCAGCCGGA GACCCCGGG CAACTCTGGC TCACAGTACG CGTACTGCAA CCGAACGCGA 2880
172 CCGCATGGTC AGAAGCCGGG CACATCAGCG CCTGGCAGCA GTGGCGTCTG GCGGAAAACC 2940
174 TCAGTGTGAC GCTCCCGGCC GCGTCCCACG CCATCCCGCA TCTGACCACC AGCGAAATGG 3000
176 ATTTTTCGAT CGAGCTGGGT AATAAGCGTT GGCAATTTAA CCGCCAGTCA GGCTTTCTTT 3060
178 CACAGATGTG GATTCGGCAT AAAAAACAAC TGCTGACGCC GCTGCGCGAT CAGTTCACCC 3120
180 GTGCACCGGT GGATAACGAC ATTGGCGTAA GTGAAGCGAC CCGCATTGAC CCTAACGCCCT 3180
182 GGGTCGAACG CTGGAAGGCG GCGGGCCATT ACCAGGCCGA AGCAGCGTTG TTGCAGTGCA 3240
184 CGGCAGATAC ACTTGTCTGT GCGGTGCTGA TTACGACCGC TCACGCGTGG CAGCATCAGG 3300
186 GGAAACCTTT ATTTATCAGC CGGAAAACCT ACCGGATTGA TGGTAGTGGT CAAATGGCGA 3360
188 TTACCGTTGA TGTGAAGTG GCGAGCGATA CACCGCATCC GCGCGCGATT GGCCTGAACT 3420
190 GCCAGCTGGC GCAGGTAGCA GAGCGGGTAA ACTGGCTCGG ATTAGGGCCG CAAGAAAACCT 3480
192 ATCCCGACCG CCTTACTGCC GCCIGTTTIG ACCGCTGGGA TCTGCCATTG TCAGACATGT 3540
194 ATACCCCGTA CGTCTCCCG AGCGAAAACG GTCGTGCGTG CGGGACGCGC GAATTGAATT 3600
196 ATGGCCACA CCAATGGGCG GCGGACTTCC AGTTCAACAT CAGCCGCTAC AGTCAACAGC 3660
198 AACTGATGGA AACCAGCCAT CGCCATCTGC TGCACGCGGA AGAAGGCACA TGGCTGAATA 3720
200 TCGACGGTTT CCATATGGGG ATTGGTGGCG ACGACTCCTG GAGCCCGTCA GTATCGGCGG 3780
202 AATTCAGCT GAGCGCGGT CGCTACCAAT ACCAGTTGGT CTGGTGTCAA AAATAATAAT 3840
204 AACCGGGCAG GCCATGTCTG CCGTATTTT CCGTAAGGAA ATCCATTATG TACTATTTAA 3900
206 AAAACACAAA CTTTGGATG TTGGTATTAT TCTTTTCTT TACTTTTTT ATCATGGGAG 3960
208 CCTACTTCCC GTTTTCCCG ATTGGCTAC ATGACATCAA CCATATCAGC AAAAGTGATA 4020
210 CCGGTATTAT TTTTCCCGT ATTTCTCTGT TCTCGCTATT ATCCAACCG CTGTTTGGTC 4080
212 TGCTTTCTGA CAAACTCGGA ACTTGTTTAT TGCAGCITAT AATGGTTACA AATAAGCAA 4140
214 TAGCATCACA AATTTACAA ATAAAGCATT TTTTCACTG CATTCAGTT GTGGTTGTGTC 4200
216 CAAACTCATC AATGTATCTT ATCATGTCTG GATCCTCTAG AGTCGACCTG CAGGCATGCA 4260
218 AGCTGGCACT GGCCGTCGT
220 (2) INFORMATION FOR SEQ ID NO: 2:
222 (i) SEQUENCE CHARACTERISTICS:
223 (A) LENGTH: 20 base pairs
224 (B) TYPE: nucleic acid
225 (C) STRANDEDNESS: single
226 (D) TOPOLOGY: linear
228 (ii) MOLECULE TYPE: other nucleic acid
229 (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
234 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
236 GAATACAAAG CTTATGCATG 20
238 (2) INFORMATION FOR SEQ ID NO: 3:
240 (i) SEQUENCE CHARACTERISTICS:
241 (A) LENGTH: 13 base pairs
242 (B) TYPE: nucleic acid
243 (C) STRANDEDNESS: single
244 (D) TOPOLOGY: linear
246 (ii) MOLECULE TYPE: other nucleic acid
247 (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
252 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

RAW SEQUENCE LISTING                      DATE: 11/29/2000  
 PATENT APPLICATION: US/09/446,402        TIME: 13:41:09

Input Set : A:\Lucyl.app  
 Output Set: N:\CRF3\11292000\I446402.raw

```

254 GAATACAAAG CTT                                     13
256 (2) INFORMATION FOR SEQ ID NO: 4:
258   (i) SEQUENCE CHARACTERISTICS:
259       (A) LENGTH: 20 base pairs
260       (B) TYPE: nucleic acid
261       (C) STRANDEDNESS: single
262       (D) TOPOLOGY: linear
264   (ii) MOLECULE TYPE: other nucleic acid
265       (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
270   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
272 AAAGCTTATG CATGCGGCCG                               20
274 (2) INFORMATION FOR SEQ ID NO: 5:
276   (i) SEQUENCE CHARACTERISTICS:
277       (A) LENGTH: 20 base pairs
278       (B) TYPE: nucleic acid
279       (C) STRANDEDNESS: single
280       (D) TOPOLOGY: linear
282   (ii) MOLECULE TYPE: other nucleic acid
283       (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
288   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
290 CGGCCGCATC TAGAGGGCCC                               20
292 (2) INFORMATION FOR SEQ ID NO: 6:
294   (i) SEQUENCE CHARACTERISTICS:
295       (A) LENGTH: 25 base pairs
296       (B) TYPE: nucleic acid
297       (C) STRANDEDNESS: single
298       (D) TOPOLOGY: linear
300   (ii) MOLECULE TYPE: other nucleic acid
301       (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
306   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
308 GCGGCCGCAT CTAGAGGGCC CGGAT                          25
310 (2) INFORMATION FOR SEQ ID NO: 7:
312   (i) SEQUENCE CHARACTERISTICS:
313       (A) LENGTH: 24 base pairs
314       (B) TYPE: nucleic acid
315       (C) STRANDEDNESS: single
316       (D) TOPOLOGY: linear
318   (ii) MOLECULE TYPE: other nucleic acid
319       (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
324   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
326 AATACAAAGC TTATGCATGC GGCC                          24
328 (2) INFORMATION FOR SEQ ID NO: 8:
330   (i) SEQUENCE CHARACTERISTICS:
331       (A) LENGTH: 30 base pairs
332       (B) TYPE: nucleic acid
333       (C) STRANDEDNESS: single
334       (D) TOPOLOGY: linear
336   (ii) MOLECULE TYPE: other nucleic acid
337       (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"

```

RAW SEQUENCE LISTING                      DATE: 11/29/2000  
 PATENT APPLICATION: US/09/446,402        TIME: 13:41:09

Input Set : A:\Lucyl.app  
 Output Set: N:\CRF3\11292000\I446402.raw

```

342      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
344 AATACAAAGC TTATGCATGC GCCCGCATCT      30
346 (2) INFORMATION FOR SEQ ID NO: 9:
348      (i) SEQUENCE CHARACTERISTICS:
349          (A) LENGTH: 20 base pairs
350          (B) TYPE: nucleic acid
351          (C) STRANDEDNESS: single
352          (D) TOPOLOGY: linear
354      (ii) MOLECULE TYPE: other nucleic acid
355          (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
360      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
362 CATGCATAAG CTTGTATTCT      20
364 (2) INFORMATION FOR SEQ ID NO: 10:
366      (i) SEQUENCE CHARACTERISTICS:
367          (A) LENGTH: 13 base pairs
368          (B) TYPE: nucleic acid
369          (C) STRANDEDNESS: single
370          (D) TOPOLOGY: linear
372      (ii) MOLECULE TYPE: other nucleic acid
373          (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
378      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
380 AAGCTTTGTA TTC      13
382 (2) INFORMATION FOR SEQ ID NO: 11:
384      (i) SEQUENCE CHARACTERISTICS:
385          (A) LENGTH: 20 base pairs
386          (B) TYPE: nucleic acid
387          (C) STRANDEDNESS: single
388          (D) TOPOLOGY: linear
390      (ii) MOLECULE TYPE: other nucleic acid
391          (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
396      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
398 CGGCCGCATG CATAAGCTTT      20
400 (2) INFORMATION FOR SEQ ID NO: 12:
402      (i) SEQUENCE CHARACTERISTICS:
403          (A) LENGTH: 20 base pairs
404          (B) TYPE: nucleic acid
405          (C) STRANDEDNESS: single
406          (D) TOPOLOGY: linear
408      (ii) MOLECULE TYPE: other nucleic acid
409          (A) DESCRIPTION: /desc = "Synthetic oligonucleotide"
414      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
416 GGGCCCTCTA GATGCGGCCG      20
418 (2) INFORMATION FOR SEQ ID NO: 13:
420      (i) SEQUENCE CHARACTERISTICS:
421          (A) LENGTH: 25 base pairs
422          (B) TYPE: nucleic acid
423          (C) STRANDEDNESS: single
424          (D) TOPOLOGY: linear
426      (ii) MOLECULE TYPE: other nucleic acid

```

VERIFICATION SUMMARY                      DATE: 11/29/2000  
PATENT APPLICATION: US/09/446,402        TIME: 13:41:10

Input Set : A:\Lucyl.app  
Output Set: N:\CRF3\11292000\I446402.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]